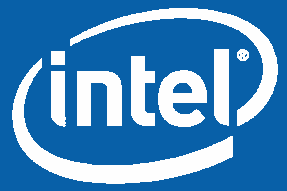


Impact of Cache Coherence Protocols on the Processing of Network Traffic

Amit Kumar and Ram Huggahalli

Communication Technology Lab
Corporate Technology Group
Intel Corporation

12/3/2007

Outline

Background

- ➔ Network performance improvement with new micro-architecture
- ➔ Need to revisit platform changes for CPU on loading

Overview of existing and Prefetch-hint coherence protocols

Direct Cache Access (DCA) Performance Overview

Prototype Results

Future Research

Background & Motivation

- Adoption of 10Gbps has been limited to a few applications. A primary reason has been the processing capability of general purpose platforms.
- Recent micro-architectural changes offered by Intel® Core™ processors has shown 66% higher network processing capability over a previous generation Intel® Pentium® 4 architecture
- Providing a coherence protocol that places data into CPU cache further improves processing capabilities
- Our prototype implementation of Direct Cache Access (DCA) shows 15.6% - 43.4% speed up

Background & Motivation (contd.)

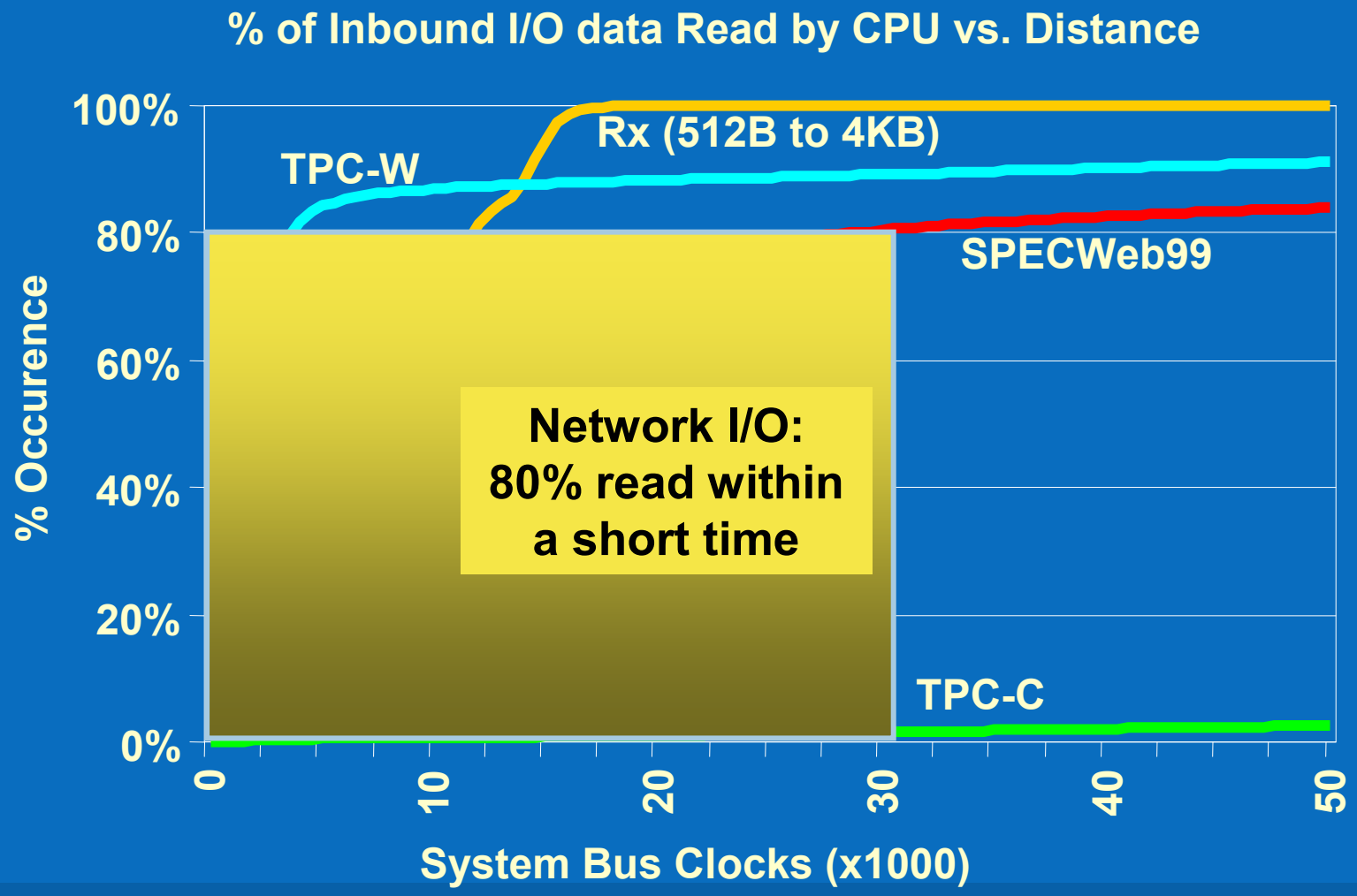
Solutions to reduce TCP/IP processing overhead can be classified in three categories:

- ➔ Platform improvements to improve CPU on loading – Copy specific solutions have been user level TCP/IP stack, Page flipping etc.
- ➔ TCP Offload Engines (TOEs) – Uses hardware assists to offload main CPU. Limited to small spectrum of networking applications.
- ➔ Interconnects or protocols like Infiniband, Myrinet or RDMA – Requires new hardware-software interfaces which requires application support. In some cases, it requires expensive NIC solutions as well.

New micro-architectural efficiencies provide a greater impetus for CPU on loading and diminishes need of specialized solutions.

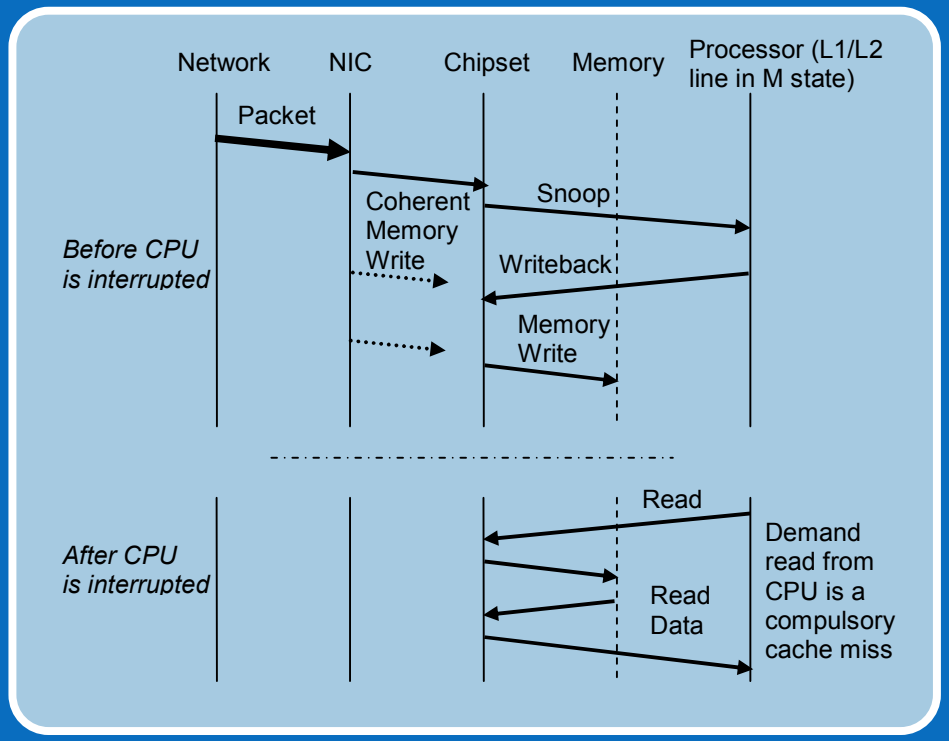
Opportunity for DCA in Realistic Workloads

Source: "Direct Cache Access for High Bandwidth Network I/O". 32nd Annual International Symposium on Computer Architecture (ISCA'05) pp. 50-59. Ram Huggahalli, Ravi Iyer and Scott Tetrick.



Today's Coherence Protocol

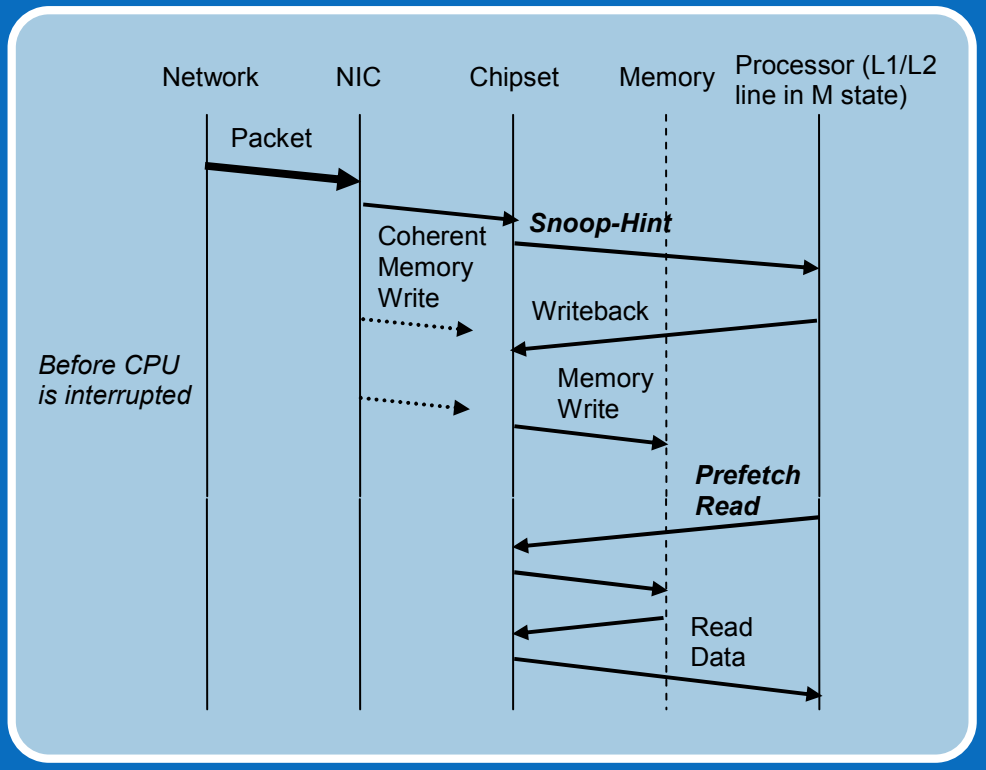
1. Packet arrives on the NIC from the network
2. NIC sends the packet as I/O bus transactions to the Chipset
3. Chipset ensures coherency of data by snooping processor caches before writing to memory
4. Processor eventually reads packet for TCP/IP processing and moves data to application buffer



Coherence protocol for inbound I/O

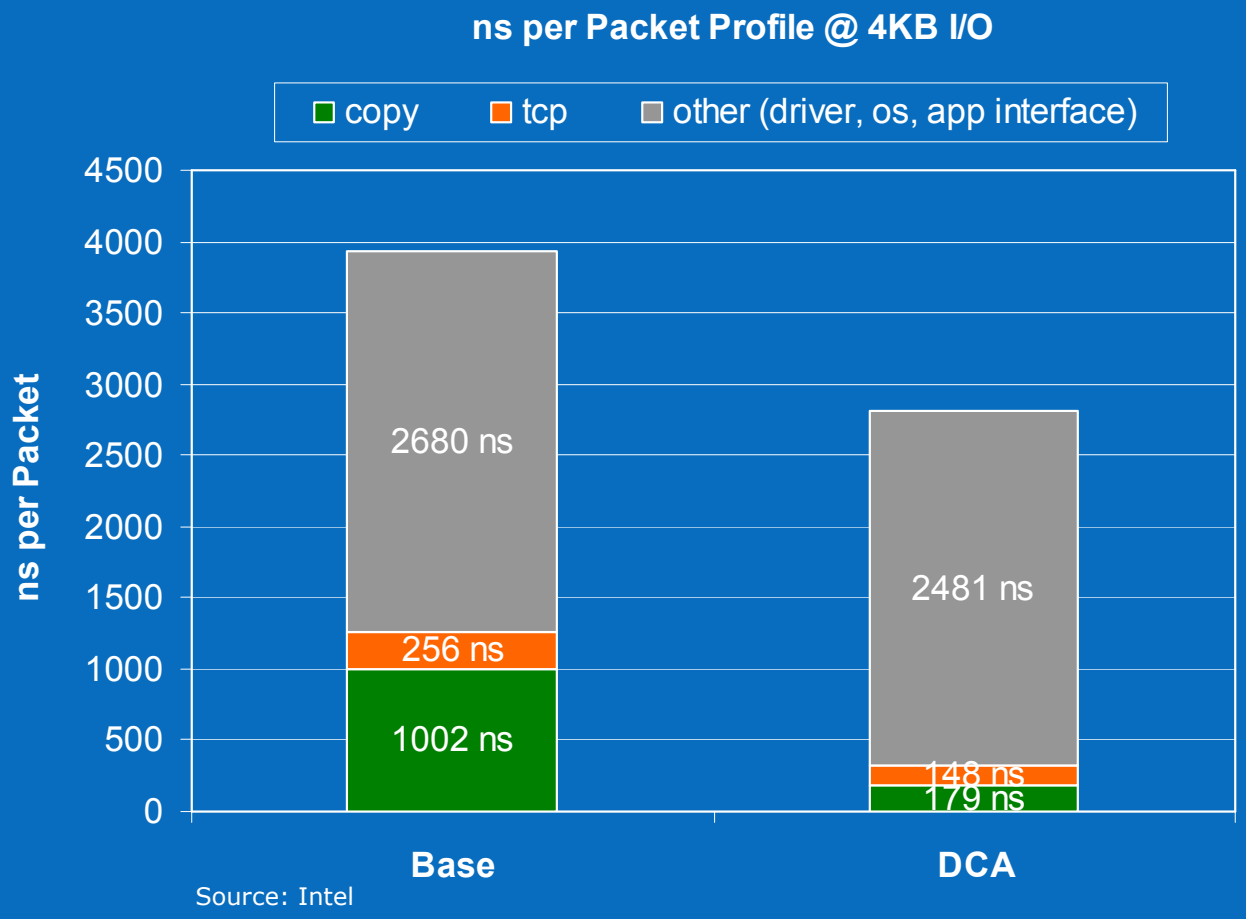
Prefetch Hint Protocol

1. Packet arrives on the NIC from the network
2. NIC sends the packet as I/O bus transactions (with a target cache tag) to the Chipset
3. Chipset sends snoops to the processor with hints to prefetch the data
4. Processor prefetches packet soon after hint is received. Packet is present in the cache TCP/IP processing begins

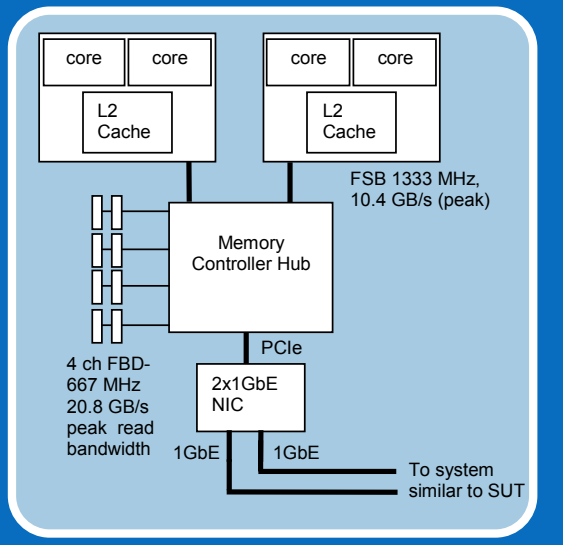


Coherence protocol for DCA prototype

Impact of Prefetch Hint/DCA protocol



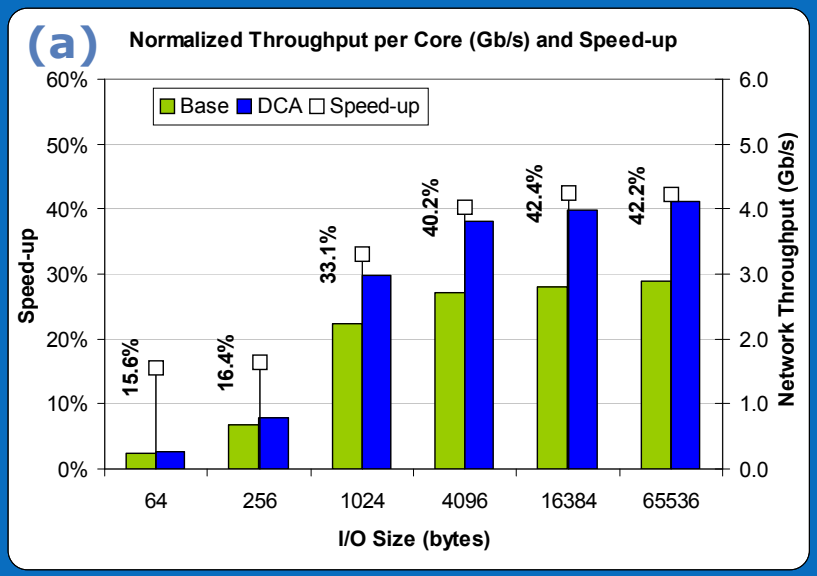
Source: Intel

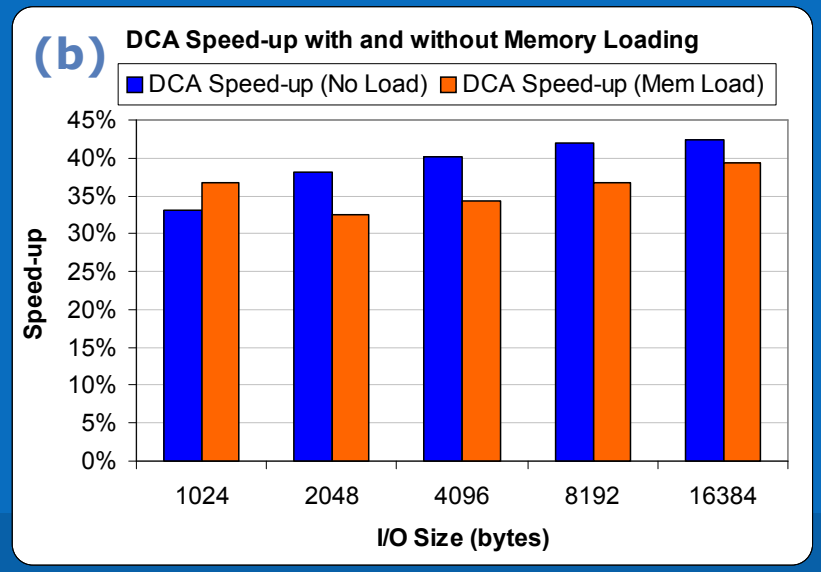


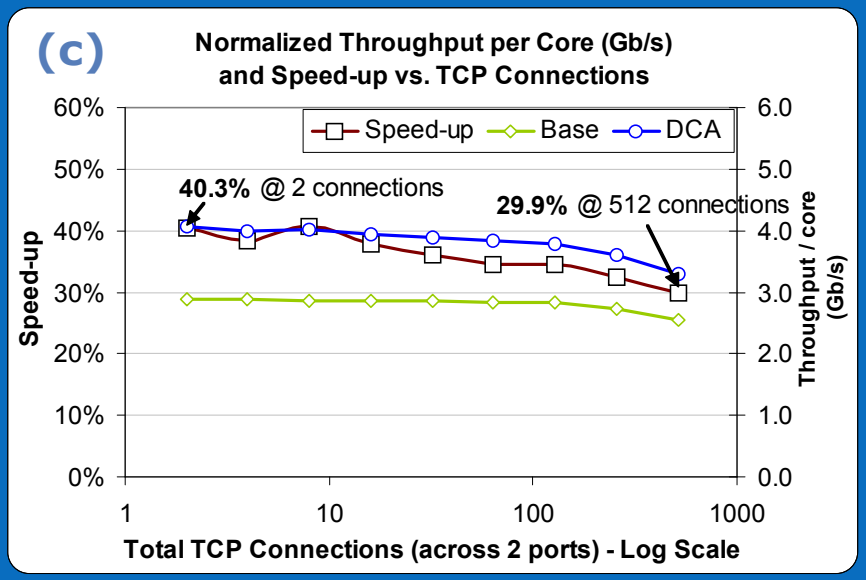
System Configuration

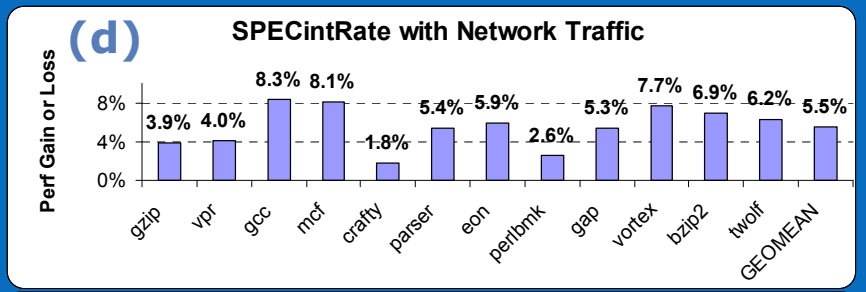
Copy with DCA is 5x faster and TCP/IP processing is 1.5x faster

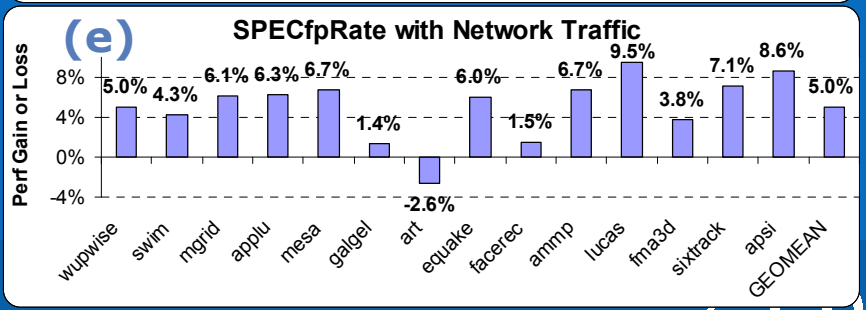
DCA Performance & Sensitivities











Future Research

DCA next steps:

- ➔ Protocol Optimization – Bypass memory and write incoming data directory into LLC (Write Update protocol)
- ➔ Performance improvement with DCA at 10Gbps and real application benefit

Related future work:

- ➔ Read Current – It is a network transmit optimization where the cached buffer used to transmit data remains in the same state in the cache
- ➔ Cache QoS – Network processing cycles through kernel buffers through the CPU cache evicting other useful data. Cache QoS policies will restrict such pollution by restricting network data to few ways in the cache
- ➔ CPU-NIC Integration – Integrating NIC on CPU can unveil many opportunities that traditional SW and HW don't enjoy. A bigger ecosystem uplift is required to make effective use of NIC integration

Q&A

Disclaimer: Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other sources of information to evaluate the performance of systems or components they are considering purchasing. For more information on performance tests and on the performance of Intel products, visit http://www.intel.com/performance/resources/limits.htm